SV.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/282,619	03/31/1999	QUAN G. CUNG	AT9-99-037	8855

7590 04/07/2004

BRACEWELL & PATTERSON LLP
INTELLECTUAL PROPERTY LAW
P.O. Box 969
Austin, TX 78767-0969

EXAMINER	
DAY, HERNG DER	
ART UNIT	**PAPER NUMBER**
2128	

DATE MAILED: 04/07/2004 22

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

Sl

UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450
www.uspto.gov

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Paper No. 22

Application Number: 09/282,619
Filing Date: March 31, 1999
Appellants: CUNG ET AL.

Andrew J. Dillon
For Appellant

MAILED
APR 0 7 2004
Technology Center 2100

EXAMINER'S ANSWER

This is in response to the appeal brief filed January 20, 2004.

(1) Real Party in Interest

A statement identifying the real party in interest is contained in the brief.

(2) Related Appeals and Interferences

A statement identifying the related appeals and interferences which will directly affect or be directly affected by or have a bearing on the decision in the pending appeal is contained in the brief.

(3) Status of Claims

The statement of the status of the claims contained in the brief is correct.

(4) Status of Amendments After Final

The Appellants' statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Invention

The summary of invention contained in the brief is deficient because the Appellants improperly introduce alleged benefits of efficiency and accuracy from this invention (e.g., first five lines in page 3 and last eight lines in page 4, Appeal Brief), which is irrelevant as it relates to "*Summary of the Invention*". Such alleged benefits have not been considered.

(6) Issues

The Appellants' statement of the issues in the brief is correct.

(7) Grouping of Claims

Appellants' statement identifies claims 1, 5-6, 13-15, 18-22 and 25 stand or fall together as a first group, claims 3-4, 16-17 and 23-24 stand or fall together as a second group.

(8) Claims Appealed

The copy of the appealed claims contained in the Appendix to the brief is correct.

(9) Prior Art of Record

5,692,107	SIMOUDIS ET AL.	11-1997

Piatetsky-Shapiro, "Discovery, Analysis, and Presentation of Strong Rules", in "Knowledge Discovery in Database", AAAI/MIT Press, 1991, pp. 229-248.

Dash et al., "Dimensionality Reduction of Unsupervised Data", Proceedings, Ninth IEEE International Conference on Tools with Artificial Intelligence, Nov. 1997, pp. 532-539.

(10) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

A. Claims 1, 5-6, 13-15, 18-22, and 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Piatetsky-Shapiro, "Discovery, Analysis, and Presentation of Strong Rules", in "Knowledge Discovery in Database", AAAI/MIT Press, 1991, in view of Simoudis et al., U.S. Patent No. 5,692,107 issued on November 25, 1997.

Piatetsky-Shapiro discloses KID3 Algorithm for discovery of exact rules. "At the end, a cell for A=a contains the summary of all the file tuples satisfying A=a" (page 235, line 34). The summary may preserve all of the field values and their relation to one another because "what intermediate information has to be kept is determined by the type of the summary description we want to have at the end" (page 236, lines 6-8). Therefore, one of ordinary skill in the art would

be able to obtain a target group with one or more desired attributes and respective values, for example, A=a and B=b, by applying KID3 Algorithm and setting the summary to collect all the information of every sample which satisfying A=a and B=b within the sample population.

Piatetsky-Shapiro also discloses how to precompute field statistics (page 230, line 19) to get information such as "How many samples satisfy the condition of C=c?" (page 233, lines 13-16) in the sample population. With these statistics, the cardinality |C| of conditions C=c can be estimated efficiently and accurately (page 233, lines 19-20).

Besides, Piatetsky-Shapiro teaches, "the typical rule-discovery task is to find K rules with the highest rule-interest function" (page 231, line 28). "Usually, the interest of rule A → B is computed as a function of p(A), the probability of A; p(B); p(A&B); rule complexity, and, possibly, other parameters, such as the mutual distribution of A and B or the domain sizes of A and B" (page 231, lines23-26). The simplest function that Piatetsky-Shapiro suggested is |A&B| - (|A||B|/N) (page 232, line 18). By using the simplest rule-interest function as Piatetsky-Shapiro suggested, a simple statistical measure will be calculated. Using a different rule-interest function, such as an entropy function, may generate a different statistical measure.

A-1. Regarding claims 1 and 5-6, Piatetsky-Shapiro discloses a method of reducing the number of the number of attributes and respective values of a sample population employed in generating a predictive model, said method comprising the steps of:

obtaining one or more desired attributes and respective values (A=a, page 235, line 3);

comparing said one or more desired attributes and respective values with said sample population to obtain a target population (KID3 Algorithm and summary, page 235, lines 21-34);

determining a statistical measure of difference (rule-interest measures, section 13.3; and the simplest function, page 232, line 18) between each of the attributes and respective values of said target population (KID3 Algorithm, page 235, lines 21-34) and the attributes and respective values of the sample population (Precomputing Field-Value Statistics, section 13.4).

Piatetsky-Shapiro fails to expressly disclose the details of utilizing the statistical measure of difference to reduce the number of attributes and respective values of said sample population, although Piatetsky-Shapiro discloses that patterns of full data set can be estimated by sample-derived rules with an estimated accuracy (Abstract), i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. Nevertheless, Piatetsky-Shapiro does suggest that rules are part of the domain knowledge (page 230, line 12) and domain knowledge can be used to improve on purely statistical prediction (page 230, line 26).

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques (column 3, lines 15-25), for example, a statistical module, to perform data mining to the selected data set to gain domain knowledge (for example, collection of rules, column 4, lines 53-55). The user sets module-specific parameters. One of ordinary skill in the art of extracting a predictive model knows that the module-specific parameters may include the criteria of extraction, which is a design choice. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

(claim 1) utilizing said statistical measure of difference to reduce the number of attributes and respective values of said sample population (a predictive model is extracted based on data mining results, column 4, lines 52-53);

(claim 5) identifying a predetermined percentage of attributes and respective values having a larger statistical measure of difference than remaining attributes and respective values (user's queries or hypotheses set module-specific parameters, column 4, lines 45-48);

(claim 6) identifying attributes and respective values where said statistical measure of difference exceeds a predetermined amount (user's queries or hypotheses set module-specific parameters, column 4, lines 45-48).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Piatetsky-Shapiro to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claims 1 and 5-6 because Simoudis et al. disclose details of a very flexible data mining method, for example, several different modules implementing different data mining techniques to choose from, for generating predictive models to gain domain knowledge as Piatetsky-Shapiro suggested to improve a purely statistical prediction (page 230, line 26).

A-2. Regarding claim 13, Piatetsky-Shapiro discloses a method of selecting attributes for computing a model, comprising:

for a plurality of samples each having values for a plurality of attributes (Extending KID3 to Complex Conditions, section 13.5.3, page 237):

for each of the plurality of attributes:

comparing the attribute values for a target group of samples to the attribute values for all of the plurality of samples (KID3 Algorithm, page 235, lines 21-34); and

determining a difference between the attribute values for the target groups and the attribute values for all of the plurality of samples (the simplest function, page 232, line 18).

Piatetsky-Shapiro fails to expressly disclose the details of identifying and selecting attributes, although Piatetsky-Shapiro discloses that patterns of full data set can be estimated by sample-derived rules with an estimated accuracy (Abstract), i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. Nevertheless, Piatetsky-Shapiro does suggest that rules are part of the domain knowledge (page 230, line 12) and domain knowledge can be used to improve on purely statistical prediction (page 230, line 26).

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques (column 3, lines 15-25), for example, a statistical module, to perform data mining to the selected data set to gain domain knowledge (for example, collection of rules, column 4, lines 53-55). The user sets module-specific parameters. One of ordinary skill in the art of extracting a predictive model knows that the module-specific parameters may include the criteria of extraction, which is a design choice. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

identifying attributes within the plurality of attributes having a largest difference between the attribute values for the target groups and the attribute values for all of the plurality of samples (user's queries or hypotheses set module-specific parameters, column 4, lines 45-48); and

selecting at least some of the identified attributes (user's queries or hypotheses set module-specific parameters, column 4, lines 45-48).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Piatetsky-Shapiro to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claim 13 because Simoudis et al. disclose details of a very flexible data mining method for generating predictive models to gain domain knowledge as Piatetsky-Shapiro suggested to improve a purely statistical prediction (page 230, line 26).

A-3. Regarding claims 14 and 18-19, Piatetsky-Shapiro discloses a system for selecting attributes for computing a model, comprising:

a memory containing data for a plurality of samples each having values for a plurality of attributes (Extending KID3 to Complex Conditions, section 13.5.3, page 237); and

a processor coupled to the memory and executing a selection process including:

comparing attribute values for samples having a desired attribute value to attribute values for all samples (KID3 Algorithm, page 235, lines 21-34);

Piatetsky-Shapiro fails to expressly disclose the details of selecting a subset of available attributes and employing the selected subset of attributes to generate a predictive model, although Piatetsky-Shapiro discloses that patterns of full data set can be estimated by sample-derived rules with an estimated accuracy (Abstract), i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. Nevertheless, Piatetsky-

Shapiro does suggest that rules are part of the domain knowledge (page 230, line 12) and domain knowledge can be used to improve on purely statistical prediction (page 230, line 26).

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques (column 3, lines 15-25), for example, a statistical module, to perform data mining to the selected data set to gain domain knowledge (for example, collection of rules, column 4, lines 53-55). The user sets module-specific parameters. One of ordinary skill in the art of extracting a predictive model knows that the module-specific parameters may include the criteria of extraction, which is a design choice. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

(claim 14) selecting a subset of available attributes (results, column 4, lines 48-51) based on a difference between attribute values for the samples having the desired attribute value and attribute values for all of the samples (Piatetsky-Shapiro, rule-interest measures, section 13.3; and the simplest function, page 232, line 18); and

(claim 14) employing the selected subset of attributes to generate a predictive model (a predictive model is extracted based on data mining results, column 4, lines 52-53);

(claim 18) identifies a predetermined percentage of attributes having a larger difference in the attribute values for selection (user's queries or hypotheses set module-specific parameters, column 4, lines 45-48);

(claim 19) identifies, for selection, attributes having a difference in the attribute values exceeding a predetermined amount (user's queries or hypotheses set module-specific parameters, column 4, lines 45-48).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Piatetsky-Shapiro to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claims 14 and 18-19 because Simoudis et al. disclose details of a very flexible data mining method for generating predictive models to gain domain knowledge as Piatetsky-Shapiro suggested to improve a purely statistical prediction (page 230, line 26).

A-4. Regarding claim 15, Piatetsky-Shapiro further discloses that the selection process determines a statistical measure of difference between the attribute values for samples having the desired attribute and the attribute values for all of the samples (the simplest function, page 232, line 18).

A-5. Regarding claim 20, Piatetsky-Shapiro discloses a system for computing a model, comprising:

a memory containing data for a plurality of samples each having values for a plurality of attributes (Extending KID3 to Complex Conditions, section 13.5.3, page 237); and

a processor coupled to the memory and executing a selection process including:

comparing attribute values for a target subset of the plurality of samples to attribute values for all of the samples (KID3 Algorithm, page 235, lines 21-34);

Piatetsky-Shapiro fails to expressly disclose the details of selecting attributes and computing a model, although Piatetsky-Shapiro discloses that patterns of full data set can be

estimated by sample-derived rules with an estimated accuracy (Abstract), i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. Nevertheless, Piatetsky-Shapiro does suggest that rules are part of the domain knowledge (page 230, line 12) and domain knowledge can be used to improve on purely statistical prediction (page 230, line 26).

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques (column 3, lines 15-25), for example, a statistical module, to perform data mining to the selected data set to gain domain knowledge (for example, collection of rules, column 4, lines 53-55). The user sets module-specific parameters. One of ordinary skill in the art of extracting a predictive model knows that the module-specific parameters may include the criteria of extraction, which is a design choice. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

selecting attributes having a largest difference between attribute values for the target subset and attribute values for all of the samples (results, column 4, lines 48-51); and

computing a model employing the selected attributes (a predictive model is extracted based on data mining results, column 4, lines 52-53).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Piatetsky-Shapiro to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claim 20 because Simoudis et al. disclose details of a

very flexible data mining method for generating predictive models to gain domain knowledge as Piatetsky-Shapiro suggested to improve a purely statistical prediction (page 230, line 26).

A-6. Regarding claims 21 and 22, Piatetsky-Shapiro discloses a computer usable medium for selecting attributes for computing a model, said computer usable medium comprising:

computer program code for reading values of attributes for a plurality of samples (Extending KID3 to Complex Conditions, section 13.5.3, page 237);

computer program code for comparing attribute values for samples having a desired attribute value to attribute values for all samples (KID3 Algorithm, page 235, lines 21-34);

Piatetsky-Shapiro fails to expressly disclose the details of selecting a subset of available attributes and determining a statistical measure of difference, although Piatetsky-Shapiro discloses that patterns of full data set can be estimated by sample-derived rules with an estimated accuracy (Abstract), i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. Nevertheless, Piatetsky-Shapiro does suggest that rules are part of the domain knowledge (page 230, line 12) and domain knowledge can be used to improve on purely statistical prediction (page 230, line 26).

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques (column 3, lines 15-25), for example, a statistical module, to perform data mining to the selected data set to gain domain knowledge (for example, collection of rules, column 4, lines 53-55). The user sets module-specific parameters. One of ordinary skill in the art of extracting a predictive model knows that the module-specific parameters may include the criteria of extraction, which is a design choice. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a

predictive model is extracted based on such results (column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

(claim 21) computer program code for selecting a subset of available attributes based on a difference between attribute values for samples having the desired attribute value and attribute values for all samples (results, column 4, lines 48-51);

(claim 22) computer program code for determining a statistical measure of difference between the attribute values for samples having the desired attribute value and the attribute values for all samples (the simplest function, page 232, line 18).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Piatetsky-Shapiro to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claims 21 and 22 because Simoudis et al. disclose details of a very flexible data mining method for generating predictive models to gain domain knowledge as Piatetsky-Shapiro suggested to improve a purely statistical prediction (page 230, line 26).

A-7. Regarding claim 25, Piatetsky-Shapiro discloses a computer usable medium for selecting attributes for computing a model, said computer usable medium comprising:

computer program code for comparing attribute values for a target group of samples to attribute values for all samples for each of a plurality of attributes (KID3 Algorithm, page 235, lines 21-34);

computer program code for determining a difference between the attribute values for the target group of samples and the attribute values for all of the samples (the simplest function, page 232, line 18); and

Piatetsky-Shapiro fails to expressly disclose the details of selecting a group of attributes, although Piatetsky-Shapiro discloses that patterns of full data set can be estimated by sample-derived rules with an estimated accuracy (Abstract), i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. Nevertheless, Piatetsky-Shapiro does suggest that rules are part of the domain knowledge (page 230, line 12) and domain knowledge can be used to improve on purely statistical prediction (page 230, line 26).

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques (column 3, lines 15-25), for example, a statistical module, to perform data mining to the selected data set to gain domain knowledge (for example, collection of rules, column 4, lines 53-55). The user sets module-specific parameters. One of ordinary skill in the art of extracting a predictive model knows that the module-specific parameters may include the criteria of extraction, which is a design choice. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

computer program code for selecting a group of attributes having a largest difference between the attribute values for the target group of samples and the attribute values for all samples (results, column 4, lines 48-51).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Piatetsky-Shapiro to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claim 25 because Simoudis et al. disclose details of a

very flexible data mining method for generating predictive models to gain domain knowledge as Piatetsky-Shapiro suggested to improve a purely statistical prediction (page 230, line 26).

B. Claims 3-4, 16-17, and 23-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over the combined teachings of Piatetsky-Shapiro, "Discovery, Analysis, and Presentation of Strong Rules", in "Knowledge Discovery in Database", AAAI/MIT Press, 1991, and Simoudis et al., U.S. Patent No. 5,692,107 issued on November 25, 1997, and further in view of Dash et al., "Dimensionality Reduction of Unsupervised Data", Proceedings, Ninth IEEE International Conference on Tools with Artificial Intelligence, Nov. 1997.

B-1. Regarding claims 3 and 4, the combined teachings of Piatetsky-Shapiro and Simoudis et al. disclose determining a statistical measure of difference in claim 1. Piatetsky-Shapiro fails to expressly disclose:

(1) determining an entropy for the attribute values;

(2) identifying n attributes having a largest difference in respective values with said target population.

Nevertheless, Piatetsky-Shapiro does suggest that in addition to the exemplary simplest function, user-specified interest weights for different fields can also be considered (page 231, lines 26-27).

Dash et al. teach an entropy measure for determining the relative importance of variables (section 2). Dash et al. also discloses a simple way to decide how many variables should be kept for a task by choosing the first d variables if it is known that an application only needs d variables (section 3). Thus the user gains insight into the data after the important original features are known. Specifically, Dash et al. disclose the missing elements:

(claim 3) determining an entropy for the attribute values (equation (1), page 534);

(claim 4) identifying n attributes having a largest difference in respective values with said target population (choose the first d variables, page 535, column 1, second paragraph).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the combined teachings of Piatetsky-Shapiro and Simoudis et al. to incorporate the teachings of Dash et al. to obtain the invention as specified in claims 3 and 4 because Dash et al. disclose details of an entropy measure which is a user-specified rule-interest measure as Piatetsky-Shapiro suggested and is applied to determine the relative importance of variables.

B-2. Regarding claim 16, the combined teachings of Piatetsky-Shapiro and Simoudis et al. disclose determining a statistical measure of difference in claim 15. Piatetsky-Shapiro fails to expressly disclose that the selection process determines an entropy for the attribute values.

Nevertheless, Piatetsky-Shapiro does suggest that in addition to the exemplary simplest function, user-specified interest weights for different fields can also be considered (page 231, lines 26-27).

Dash et al. teach an entropy measure for determining the relative importance of variables (section 2). Dash et al. also discloses a simple way to decide how many variables should be kept for a task by choosing the first d variables if it is known that an application only needs d variables (section 3). Thus the user gains insight into the data after the important original features are known. Specifically, Dash et al. disclose the missing element:

determines an entropy for the attribute values (equation (1), page 534);

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the combined teachings of Piatetsky-Shapiro and Simoudis et al. to incorporate the teachings of Dash et al. to obtain the invention as specified in claim 16 because Dash et al. disclose details of an entropy measure which is a user-specified rule-interest measure as Piatetsky-Shapiro suggested and is applied to determine the relative importance of variables.

B-3. Regarding claim 17, the combined teachings of Piatetsky-Shapiro and Simoudis et al. disclose determining a statistical measure of difference in claim 14. Piatetsky-Shapiro fails to expressly disclose that the selection process identifies a predetermined number of attributes having a largest difference in the attribute values for selection.

Nevertheless, Piatetsky-Shapiro does suggest that in addition to the exemplary simplest function, user-specified interest weights for different fields can also be considered (page 231, lines 26-27).

Dash et al. teach an entropy measure for determining the relative importance of variables (section 2, page 533, column 1, last paragraph). Dash et al. also discloses a simple way to decide how many variables should be kept for a task by choosing the first d variables, from an ordered list of variables, if it is known that an application only needs d variables (section 3, page 535, column 1, first and second paragraphs), i.e., a predetermined number d. Thus, the user gains insight into the data after the important original features are known. Specifically, Dash et al. disclose the missing element:

identifies a predetermined number of attributes having a largest difference in the attribute values for selection (choose the first d variables, page 535, column 1, second paragraph).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the combined teachings of Piatetsky-Shapiro and Simoudis et al. to incorporate the teachings of Dash et al. to obtain the invention as specified in claim 17 because Dash et al. disclose details of an entropy measure which is a user-specified rule-interest measure as Piatetsky-Shapiro suggested and is applied to determine the relative importance of variables.

B-4. Regarding claim 23, the combined teachings of Piatetsky-Shapiro and Simoudis et al. disclose determining a statistical measure of difference in claim 22. Piatetsky-Shapiro fails to expressly disclose the determining and comparing entropy and comparing the relative measure of difference.

Nevertheless, Piatetsky-Shapiro does suggest that in addition to the exemplary simplest function, user-specified interest weights for different fields can also be considered (page 231, lines 26-27).

Dash et al. teach an entropy measure for determining the relative importance of variables (section 2). Dash et al. also discloses a simple way to decide how many variables should be kept for a task by choosing the first d variables if it is known that an application only needs d variables (section 3). Thus the user gains insight into the data after the important original features are known. Specifically, Dash et al. disclose the missing elements:

computer program code for determining an entropy of the attribute values for samples having the desired attribute value and an entropy of the attribute values for all samples (equation (1), page 534);

computer program code for comparing the entropy of the attribute values for samples having the desired attribute value to the entropy of the attribute values for all samples for each

attribute to determine a relative measure of difference (ordered list, page 535, column 1, first paragraph); and

computer program code for comparing the relative measure of difference of all attributes (ordered list, page 535, column 1, first paragraph).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the combined teachings of Piatetsky-Shapiro and Simoudis et al. to incorporate the teachings of Dash et al. to obtain the invention as specified in claim 23 because Dash et al. disclose details of an entropy measure which is a user-specified rule-interest measure as Piatetsky-Shapiro suggested and is applied to determine the relative importance of variables.

B-5. Regarding claim 24, the combination of Piatetsky-Shapiro and Simoudis et al. disclose determining a statistical measure of difference in claim 21. Piatetsky-Shapiro fails to expressly disclose the identifying n attributes having a largest difference in the attribute values.

Nevertheless, Piatetsky-Shapiro does suggest that in addition to the exemplary simplest function, user-specified interest weights for different fields can also be considered (page 231, lines 26-27).

Dash et al. teach an entropy measure for determining the relative importance of variables (section 2). Dash et al. also discloses a simple way to decide how many variables should be kept for a task by choosing the first d variables if it is known that an application only needs d variables (section 3). Thus the user gains insight into the data after the important original features are known. Specifically, Dash et al. disclose the missing element:

computer program code for identifying n attributes having a largest difference in the attribute values (choose the first d variables, page 535, column 1, second paragraph).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the combined teachings of Piatetsky-Shapiro and Simoudis et al. to incorporate the teachings of Dash et al. to obtain the invention as specified in claim 24 because Dash et al. disclose details of an entropy measure which is a user-specified rule-interest measure as Piatetsky-Shapiro suggested and is applied to determine the relative importance of variables.

(11) Response to Argument

Before responds to Appellants' argument, the Examiner explains what information and knowledge Piatetsky-Shapiro has disclosed based on the example disclosed at pages 9-10 of the specification. For simplicity, assume no sample has the same attribute value and A, B, and C are the only attributes of each sample.

Based on the specification, A may have three possible values: Y, N, and UNKNOWN. B has five possible values (b_1, b_2, b_3, b_4, b_5) within sample 204, but only two of those values (b_1, b_2) are found within target group 210. C also has five possible values (c_1, c_2, c_3, c_4, c_5) within sample 204, but only four of those values (c_1, c_2, c_3, c_4) are found within target group 210. Accordingly, the total number of samples, i.e., sample population, is N = 3 x 5 x 5 = 75.

As described in lines 21-23, page 9 of the specification, "Samples having the value Y for attribute A are therefore categorized as target group 210". After applying Piatetsky-Shapiro's KID3 algorithm, at the end, a cell for A = Y contains all the samples satisfying A = Y is obtained by setting the summary description to collect all the information of every sample satisfying A = Y. Accordingly, the target group 210, i.e., target population, has the number of 1 x 2 x 4 = 8 samples with A = Y, which will be calculated by Piatetsky-Shapiro's teaching of precomputing field statistics below.

To find out which attribute value of B and C is more sensitive for the purpose of generating a predictive model, one of ordinary skill in the art would be motivated by applying Piatetsky-Shapiro's algorithm for the following exemplary rules to determine:

(i). $((A = Y) \& (B = b_1)) \rightarrow (B = b_1)$; in other words, whether all samples satisfying the condition $(A = Y) \& (B = b_1)$, i.e., value b_1 of attribute B of target population, also satisfy condition $(B = b_1)$, i.e., value b_1 of attribute B of sample population.

(ii). $((A = Y) \& (C = c_1)) \rightarrow (C = c_1)$; in other words, whether all samples satisfying the condition $(A = Y) \& (C = c_1)$, i.e., value c_1 of attribute C of target population, also satisfy condition $(C = c_1)$, i.e., value c_1 of attribute C of sample population.

Next, following Piatetsky-Shapiro's teaching of precomputing field statistics to get information such as "How many samples satisfy the condition of $A = Y$ or $B = b_1$?" in the sample population. Let $|(A = Y)|$ denote the number of samples that satisfy the condition $(A = Y)$, we have:

(i). $|(A = Y)| = 8$, $|(A = Y) \& (B = b_1)| = 4$, $|(B = b_1)| = 3 \times 5 = 15$

(ii). $|(A = Y) \& (C = c_1)| = 2$, $|(C = c_1)| = 3 \times 5 = 15$

Finally, by applying the simplest function $|A\&B| - (|A||B|/N)$ as suggested by Piatetsky-Shapiro, a statistical measure of difference can be determined and compared between target population and sample population for specified attribute value.

(i). $4 - (4 \times 15 / 75) = 240 / 75 = 3.2$

(ii). $2 - (2 \times 15 / 75) = 120 / 75 = 1.6$

A. Issue 1, first group of claims 1, 5-6, 13-15, 18-22 and 25, stand or fall together.

A-1. Appellants argue, "the reference is devoid of any teaching for obtaining a target population" (page 6, second paragraph, lines 3-4, Appeal Brief) and "the result of the KID3 algorithm taught by Piatetsky-Shapiro produces a summary of the existing sample population, and not obtaining a target population" (page 6, fifth paragraph, lines 1-3, Appeal Brief). The Examiner respectfully disagrees with the Appellants' argument because the Appellants' argument appears to ignore the teachings of Piatetsky-Shapiro that the summary may preserve all of the field values and their relation to one another because "what intermediate information has to be kept is determined by the type of the summary description we want to have at the end" (Piatetsky-Shapiro, page 236, first paragraph, last three lines).

First, as described in lines 7-9, page 9 of the specification, "Data elements within sample 204 having the *desired attribute value or values* are categorized in target group 210". Also, as described in the last second and third lines, page 7 of the Appeal Brief, "selecting a target group by comparing attributes values of the sample population to *desired values*". Accordingly, the "target population", as defined in this instant application, is "all samples having the *desired attribute value or values* selected from the sample population".

Piatetsky-Shapiro discloses a KID3 Algorithm for discovery of exact rules. Specifically, "At the end, a cell for A = a contains the summary of all the file tuples satisfying A = a" (Piatetsky-Shapiro, page 235, last second line). In other words, the "cell for A = a" meets the definition of "target population" because it contains the summary of all the file tuples satisfying attribute A has the *desired value* a.

Furthermore, Piatetsky-Shapiro discloses the summary may preserve all of the field values and their relation to one another because "what intermediate information has to be kept is

determined by the type of the summary description we want to have at the end" (Piatetsky-Shapiro, page 236, first paragraph, last three lines). In other words, Piatetsky-Shapiro has taught one of ordinary skill in the art how to obtain a target population with one or more desired attributes and respective values, for example, A = a and B = b, by applying KID3 Algorithm and setting the summary to collect all the information of every sample which satisfying A = a and B = b within the sample population.

Also notes, in the exemplary claim 14 of the first group (Appellants indicate claims 1, 5-6, 13-15, 18-22 and 25 stand or fall together as a first group in GROUPING OF THE CLAIMS at page 5 of Appeal Brief), "target population" has not been expressly recited.

A-2. Appellants argue, "Piatetsky-Shapiro does not teach or suggest determining a statistical measure of difference between each of the attributes and respective values of the target population and sample population as recited in claim 1" (page 7, first paragraph, lines 1-3, Appeal Brief). The Examiner respectfully disagrees with the Appellants' argument because the Appellants' argument appears to ignore the teachings of Piatetsky-Shapiro that the rule-discovery task is to find K rules with the highest rule-interest function (Piatetsky-Shapiro, page 231, last paragraph, first line).

Usually, as disclosed by Piatetsky-Shapiro, the interest of rule A → B is computed as a function of p(A), the probability of A; p(B); p(A&B); and other parameters such as the domain sizes of A and B (Piatetsky-Shapiro, page 231, last second paragraph, lines 3-6). The simplest function that Piatetsky-Shapiro suggested is |A&B| - (|A||B|/N) (Piatetsky-Shapiro, page 232, last second paragraph, lines 3-6). Piatetsky-Shapiro also discloses how to precompute field statistics (Piatetsky-Shapiro, page 230, fifth paragraph, line 2) to get information such as "How many

samples satisfy the condition of C = c?" (Piatetsky-Shapiro, page 233, third paragraph) in the sample population. With these statistics, the cardinality |C| of conditions C = c can be estimated efficiently and accurately (Piatetsky-Shapiro, page 233, fourth paragraph, lines 3-4). In other words, Piatetsky-Shapiro does teach determining a statistical measure of difference between each of the attributes and respective values of the target population and sample population by applying the simplest function |A&B| - (|A||B|/N).

Also notes, in the exemplary claim 14 of the first group (Appellants indicate claims 1, 5-6, 13-15, 18-22 and 25 stand or fall together as a first group in GROUPING OF THE CLAIMS at page 5 of Appeal Brief), "statistical measure" has not been expressly recited.

A-3. Appellants argue, "Appellants acknowledge that Simoudis et al. teaches the use of a target population that is employed in generating a predictive model. However, Simoudis et al. does not teach "comparing said one or more desired attributes and respective values with said sample population to obtain a target population" as recited by the claims" (page 7, second paragraph, lines 2-6, Appeal Brief). The Appellants' acknowledgement is correct. However, the Appellants' argument, i.e., Simoudis et al. do not teach expressly obtaining a target population by comparing attribute values is moot because "comparing said one or more desired attributes and respective values with said sample population to obtain a target population" has been taught by Piatetsky-Shapiro.

Also notes, in the exemplary claim 14 of the first group (Appellants indicate claims 1, 5-6, 13-15, 18-22 and 25 stand or fall together as a first group in GROUPING OF THE CLAIMS at page 5 of Appeal Brief), "to obtain a target population" has not been expressly recited.

B. Issue 2, second group of claims 3-4, 16-17 and 23-24, stand or fall together.

B-1. Appellants argue, "Dash et al. are entirely silent on the subject of the reduction of variables based on a difference between the attributes and the respective values of a target group and sample population" and "these combined citations lack any teachings or suggestions of a determination of a statistical measure of difference between the attributes and the respective values of a target population and a sample population or the comparing of attributes and respective values with a sample population to obtain a target population" (page 8, second paragraph, lines 4-6 and 9-13, Appeal Brief). The Appellants' argument is moot because the Examiner has not applied the teachings of Dash et al. in the reduction of variables based on a difference between the attributes and the respective values of a target group and sample population. Simoudis et al. has already taught the reduction of variables and Piatetsky-Shapiro has already taught determining a difference between the target population and sample population for specified attribute value.

Also notes, in the exemplary claim 17 of the second group (Appellants indicate claims 3-4, 16-17 and 23-24 stand or fall together as a second group in GROUPING OF THE CLAIMS at page 5 of Appeal Brief), "statistical measure" and "obtain a target population" have not been expressly recited.

C. In summary, Examiner believes that all the claimed inventions have been disclosed by the combined teachings of Piatetsky-Shapiro, Simoudis et al., and Dash et al.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

Herng-der Day
April 5, 2004

Conferees
Kevin Teska
Hugh Jones

ANDREW J DILLON
FELSMAN BRADLEY GUNTER & DILLON
LAKEWOOD ON THE PARK SUITE 350
7600 E NORTH CAPITAL OF TEXAS HIGHWAY
AUSTIN, TX 78731

KEVIN J. TESKA
SUPERVISORY
PATENT EXAMINER